

Information System Model and Architecture Generator

by K. P. Hein

The advent of integrated, shared-data systems has made it increasingly necessary to address the application development process from the architectural and manufacturing perspective rather than from a build-as-you-go job shop viewpoint. Although the Business Systems Planning (BSP) methodology provides an enterprise-wide strategic Information Systems plan, it is still at an abstraction level that leaves the traditional gap between "requirements" and implementations untouched. The Information System Model and Architecture Generator (ISMOD) tool complements and enhances BSP by mechanizing the planning process, thus providing a facility to narrow this gap by allowing orderly and consistent top-to-bottom architectural decomposition of the enterprise environment. It is an enterprise planning vehicle and not an implementation system, but it is the first critical component to support an integrated systems architecture effort. It automates and, to a large extent, formalizes a laborious requirements documentation process preceding code development, and it does this "top to bottom," from a global, enterprise-wide, information requirements viewpoint. This paper discusses the overall architectural concepts of integrated data systems development, the place of ISMOD within it, and the specific facilities, techniques, and information provided by the system.

The Information System Model and Architecture Generator (ISMOD) was developed to aid enterprise executives and integrated data systems architects in the analysis of the information needs of an enterprise. Its primary purpose is the identification of the enterprise processes (functions) and data required to manage the resources of an enterprise. On the basis of data usage in the enterprise, the system will allow knowledgeable information systems architects to group various processes into process clusters called subsystems. Through a simulation capability, it is possible to project the impact a newly implemented subsystem may be able to achieve in terms of user satisfaction.

Traditionally, the data processing industry has developed applications by targeting operational, high-volume systems, with the primary concern being efficient machine utilization. A very specific and rigid development methodology has been devised which is very successful in implementing these "bottom-up, stand-alone" systems.

However, with the emergence of data base technology, many organizations have shifted their emphasis from stand-alone operational systems to integrated, shared-data systems. In this environment, it is extremely important that applications be viewed in the context of the whole enterprise. As a result, the development methodology used to create stand-alone applications is not appropriate, and in many cases, it is counterproductive.

The development methodology for shared-data systems must, as its first step, have a much broader view of the enterprise than was previously required. Before implementation of a particular enterprise information subsystem can be attempted, the architect must understand the whole so that during detailed design and implementation of a subsystem the interfaces to other subsystems are considered.

Various efforts, such as Business Systems Planning (BSP), have been undertaken to gain a comprehensive overview of an enterprise. However, they typically do not analyze the enterprise in sufficient detail to

© Copyright 1985 by International Business Machines Corporation. Copying in printed form for private use is permitted without payment of royalty provided that (1) each reproduction is done without alteration and (2) the *Journal* reference and IBM copyright notice are included on the first page. The title and abstract, but no other portions, of this paper may be copied or distributed royalty free without further permission by computer-based and other information-service systems. Permission to *republish* any other portion of this paper must be obtained from the Editor.

provide follow-on designers and implementors enough information to bring about the necessary foundation architecture required to meet the identified needs of such studies. Therefore, a great deal of additional work needs to be accomplished in order for the implementation team to meet its objectives. The reason why the initial study team is unable to provide sufficient detail, as part of an enterprise

All information systems are interdependent to various degrees.

analysis, is that the amount of data to be analyzed by the team is too voluminous. Inclusion of such detail would make the initial enterprise study too expensive, confusing, and time-consuming.

Utilizing the ISMOD tool and established BSP techniques, the study team can accomplish its mission in a relatively short time frame. At the same time, the study provides the required depth and level of detail for thorough enterprise analysis and support for follow-on implementation efforts.

Enterprise Information Systems architecture

In order to implement the integrated data systems, the data processing professionals need to approach the process in the same manner as the manufacturer of a product. However, the data processing department is not the creator or designer of any information system which has been developed to help manage the enterprise. Instead, the information system specialists are the automators of such systems. The system is initially conceived and designed by the manufacturing, financial, or administrative experts of the enterprise.

All information systems, whether manual or mechanized, are interdependent to various degrees. Historically, data processing implementations have disregarded this interdependency, resulting in systems which have integrity, maintenance, duplicate processing, and storage costs associated with them. This

approach to system implementation has been referred to as the traditional, or "bottom-up," methodology.

In recent years, leading thinkers, organizations, and information systems architects involved with enterprise analysis and data processing, such as Yourdon, Holland, Orr, DeMarco, and Zachman, have proclaimed that this traditional approach should be replaced with the more comprehensive "top-down" concept. This methodology advocates that the enterprise needs to be supported by information systems which are highly integrated through common sharing of data. The benefits of such systems are that data are acquired only once in the enterprise and distributed in parallel, information has a higher level of integrity for decision-making since multiple sources for the same data are eliminated, and data and maintenance costs can be substantially reduced.

Therefore, a need exists to supply the information system architect and developer with computer-aided facilities which allow him to realize the goal in a relatively reasonable time frame with reasonable resource expenditure.

ISMOD

ISMOD is designed as an aid to enterprises which desire to implement these more sophisticated integrated, shared-data systems. The "top-down" approach is used to define strategic, functional, and operational level architectures. Its analytical components are the flow of information within the enterprise in terms of using and originating processes/subsystems and data views.

A satisfaction index is used to measure the current value of the information system and to project the future value of planned information system efforts through simulation.

Enterprise processes are clustered into subsystems based on shared data to provide a boundary for implementation projects within the framework of an overall architecture.

Business Systems Planning, Application Transfer Team (ATT) Studies, and Information Quality Analysis (IQA) are some of the IBM-provided planning and study methodologies which have taken advantage of the capabilities of the system. A study can be conducted for many reasons. It can be done as an initial enterprise-wide strategic planning study, as a follow-

on subsystem data flow study, or as a validation study after part of the original study recommendations have been implemented.

Top management support. The primary objective of the system is to aid in accomplishing the integration and management of data as an enterprise resource.

**It is extremely important that a study
be supported by the top
management of the enterprise.**

As such, information systems for the enterprise will be identified which cross organizational boundaries and whose architectures are designed on the basis of the data used by the enterprise processes, rather than according to organizational or political considerations.

Therefore, it is extremely important that a study be supported by the top management of the enterprise, not only during the initial study but also during the later implementation phases. Without complete support from top management, the study team will be seriously hampered, and cooperation from user departments may be lacking to the extent that the study effort, no matter how well-conceived and executed, will be a failure.

User involvement. Involvement in the study by the business professionals of an enterprise is of absolute importance. Because of the large scope of a high-level enterprise analysis, it is essential that only individuals who have the broadest possible background in the study area be placed on the team. Users with expertise in details can be involved in the study through the interview process.

During the company analysis phase of the study, the study team will identify interviewees who can contribute significantly to the overall understanding of the mission and operation of the enterprise. Identification of interviewees who have more detailed knowledge of the enterprise should be as complete as possible so that the study has the broadest possible

input and, therefore, will receive the greatest possible support during implementation.

The data model

The system uses a data model of ten dimensions (entities) connected via certain predefined relationships. Dimensions are used to contain data collected through an interviewing process for later analysis. The dimensions discussed here are those provided with the release of the system. However, they can be redefined by the user to represent any information that lends itself to the type of analysis provided by the program. The data model is shown in Figure 1.

Organization. An organization represents any person or group of persons in the business who is responsible for and performs business processes.

Location/node. The location dimension is used to represent various physical or logical places where the enterprise houses organizations, performs processes, and requires the use of data. A node is a grouping of locations.

Process/subsystem. An enterprise process is a set of tasks and procedures which uses input data to create an information output which is in turn used by another process as an input. Processes are grouped into subsystems.

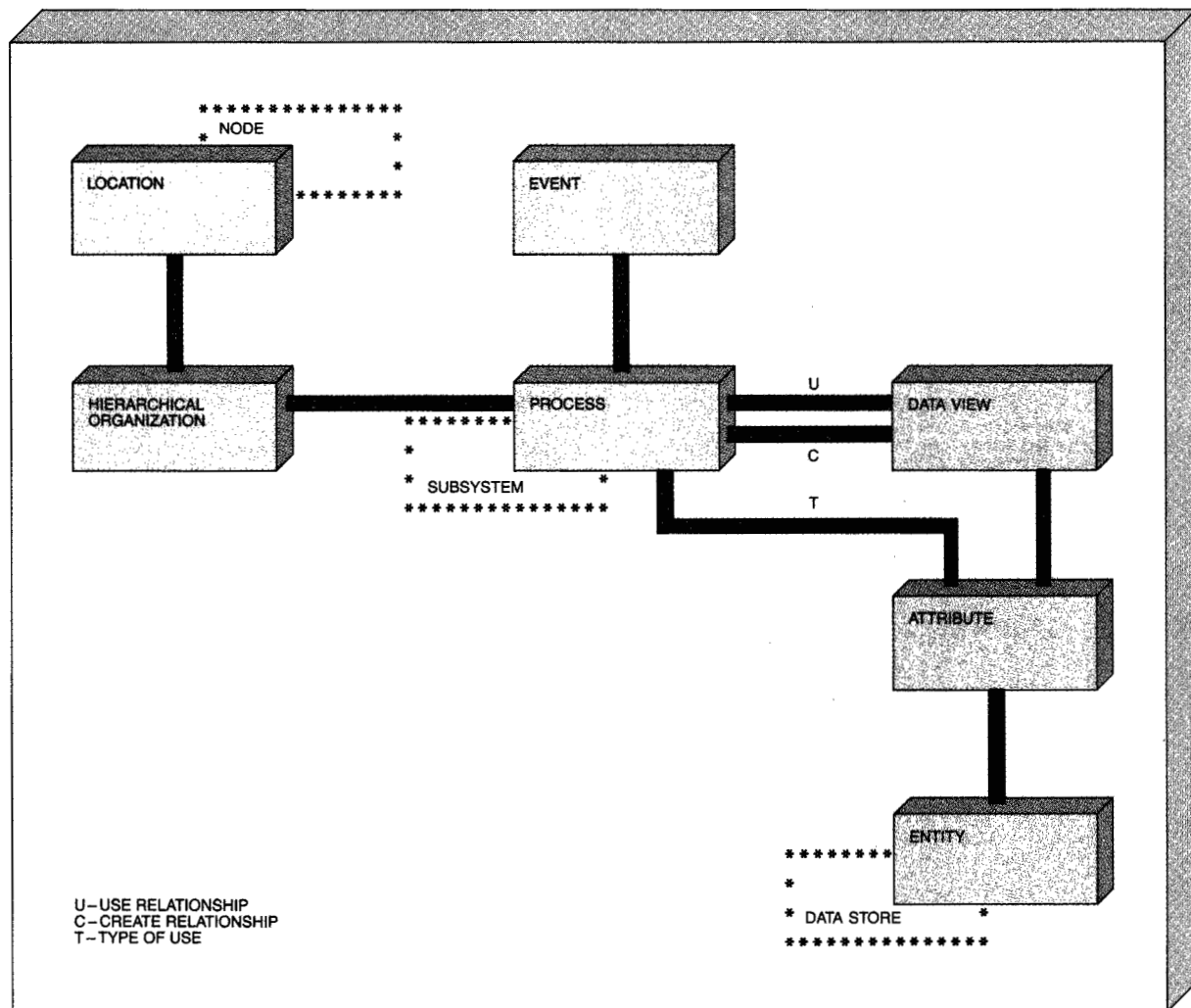
Event. The event is an occurrence in the real world which can trigger the execution of a process. For example, the "end-of-the-month" event can cause various billing processes to be performed. Other examples may be the completion of a purchase order, a customer phone call, or reaching of the reorder point.

Data view. A data view is a physical or logical grouping of elementary data into a form that allows humans to draw some logical conclusion. Examples may be an invoice, telephone message, or pay check.

Attribute. An attribute represents a specific characteristic of a person, place, thing, or idea about which the enterprise needs to maintain some information. Examples are part number, employee name, part unit of measure, and employee marital status.

Entity. This dimension represents a person, place, thing, or idea about which the business needs to keep information. Customer, employee, product, machine, and warehouse are examples of entities.

Figure 1 ISMOD model



Data store. A data store is any device that is used to house information either temporarily or permanently, such as file cabinets, in/out baskets, and disk drives.

Study phases

Enterprise planning methodologies utilizing ISMOD are generally conducted in four study phases spread over a five-to-eight-week time frame, as are depicted in Figure 2. However, these phases may be modified by various program products that use the system to support their unique advantages. The following description is one example of such a methodology.

Company analysis. During this phase, the study team gets organized and establishes a preliminary set of enterprise processes and data views/entities and an interviewee list. Additionally, the team becomes familiar with the critical success factors of the enterprise and the resources to be managed. This work effort is usually accomplished in the first week. At the end of this phase, a meeting is held with all interviewees during which they are briefed on the purpose of the upcoming interviews, how they will be conducted, and what information will be required. Prior to this briefing, results from the company analysis phase are provided to the interviewees in preparation for the briefing and interview.

Survey. The survey phase consists of three parts. First, the interviews are conducted individually or in groups. Individual interviews may be conducted at the top management level, with group interviews used from the functional management level down. Each interviewee completes a set of interview forms which describe the processes and the data required to perform the various processes.

Second, the interview forms are entered into the system by a data entry operator using system-supplied input screens.

Third, a printed copy of the interview is sent to the interviewee for validation. When the interviewee returns the corrected and updated interview summaries, they are used to update the information in the system files.

Diagnosis. During the diagnosis phase, the study team will analyze the responses of the interviewees in terms of the current, as well as future, information system needs and the levels of satisfaction associated with the required data.

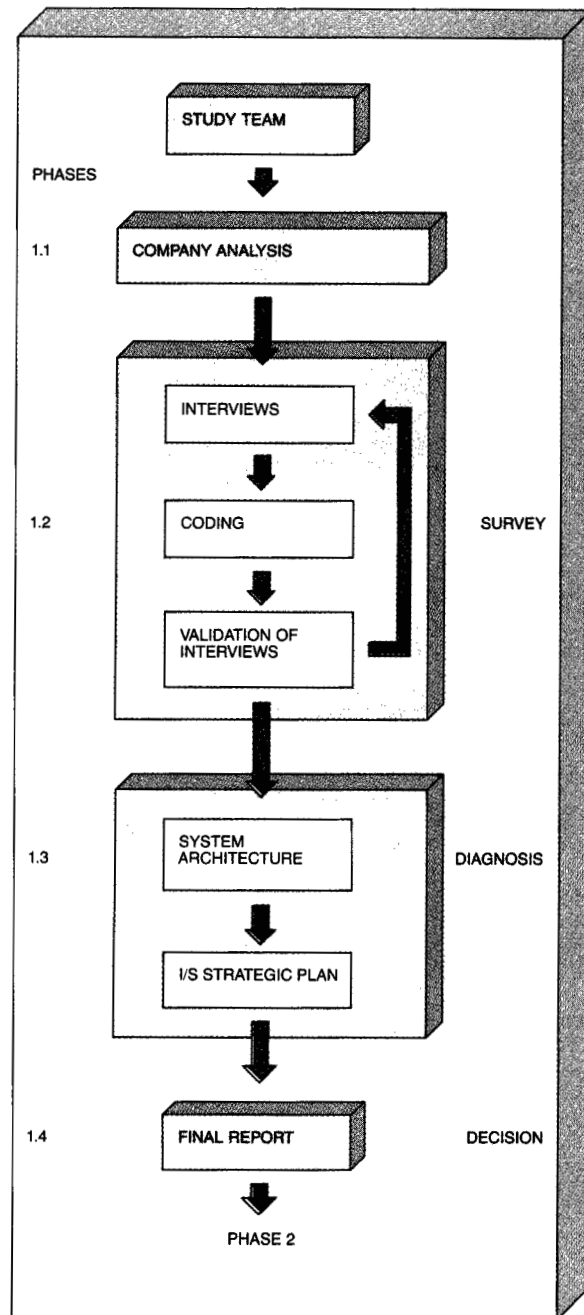
Following that, the processes are grouped into subsystems based on the amount of data interaction and sharing. After grouping has been accomplished, the study team will have a good idea of the overall enterprise Information Systems (I/S) architecture. This architecture can then be simulated for satisfaction levels if these new groups of processes (subsystems) are to be implemented. From the system architecture it is then possible to develop the I/S strategic plan.

Final report. The study team prepares a final report which will outline the proposed set of subsystems, an implementation action plan, possibly a financial cost-benefit analysis, and a recommendation for the first subsystem to be implemented. Thus, a decision is made.

Interview data

To design new information systems for the future, it is essential to understand the present I/S environment. The team must understand who is doing what, where it is being done, and what data are being used to do it. Furthermore, they must know how satisfied the organization is with various parts of the current information system. To this end, data are collected from a predetermined group of interviewees. Figures 3, 4, and 5 show examples of interview forms.

Figure 2 Study phases



Data are collected about the interviewee, his/her organization, and the location in which the inter-

Figure 3 Identification form

IBM	ISMOD	X X X	INTERVIEW IDENTIFICATION	1 3 2	PAGE	1
DATE MAY / 21 / 84			(TOTAL NUMBER OF PAGES.....)			
INFORMATION SYSTEM MODEL AND ARCHITECTURE GENERATOR						
INTERVIEW FILE						
1. IDENTIFICATION						
-INTERVIEWEE'S CODE			4 2 2			
-NAME			JOHN BROWN			
-TITLE			INVENTORY MANAGER			
-LOCATION			23 PLANT 3			

Figure 4 Process form

IBM ISMOD ☒ ☒ ☒ INTERVIEW IDENTIFICATION 1 3 2			PAGE 2
2. PROCESSES			
IDENT. 1	CODE 3 2 1	EVENT 8 3	BEGIN WEEK
MAKE PURCHASING FORECASTS			
IDENT. 2	CODE 3 2 2	EVENT 9 5	END MONTH
CONTROL STOCK MOVEMENTS			
IDENT. 3	CODE 3 2 2	EVENT 9 8	END QUARTER
CALCULATE IDEAL ORDER-POINTS			
IDENT. 4	CODE 3 2 3	EVENT 9 8	END QUARTER
OPTIMIZE SPACE IN THE WAREHOUSES			
IDENT. 5	CODE 5 1	EVENT 9 8	END QUARTER
MAKE MANPOWER PLANS			
IDENT. 6	CODE 5 3	EVENT 9 5	END MONTH
EDUCATE THE STOCK-KEEPERS			
IDENT. 7	CODE 5 4	EVENT 8 3	WEEKLY
MANAGE THE STOCK-KEEPERS			
IDENT. 8	CODE 6 1 1	EVENT 	YEARLY
ESTABLISH THE BUDGETS			

Figure 5 Interview form - data

IBM	ISMOD <table border="1" style="display: inline-table; vertical-align: middle;"><tr><td>X</td><td>X</td><td>X</td></tr></table>	X	X	X	INTERVIEW IDENTIFICATION <table border="1" style="display: inline-table; vertical-align: middle;"><tr><td>1</td><td>3</td><td>2</td></tr></table>	1	3	2	PAGE <table border="1" style="display: inline-table; vertical-align: middle;"><tr><td> </td><td> </td></tr></table>		
X	X	X									
1	3	2									

3. DATA																																																																																												
IDENT. <table border="1" style="display: inline-table;"><tr><td>1</td><td>0</td><td>1</td></tr></table>	1	0	1	CODE <table border="1" style="display: inline-table;"><tr><td> </td><td> </td><td> </td><td> </td><td> </td><td> </td></tr></table>							IDENT. OF THE PROCESSES IMPLIED																																																																																	
1	0	1																																																																																										
		<table border="1" style="display: inline-table;"><tr><td> </td><td>1</td><td> </td></tr><tr><td> </td><td> </td><td> </td></tr><tr><td> </td><td> </td><td> </td></tr><tr><td> </td><td> </td><td> </td></tr></table>		1											<table border="1" style="display: inline-table;"><tr><td> </td><td>3</td><td> </td></tr><tr><td> </td><td> </td><td> </td></tr><tr><td> </td><td> </td><td> </td></tr><tr><td> </td><td> </td><td> </td></tr></table>		3											<table border="1" style="display: inline-table;"><tr><td> </td><td>4</td><td> </td></tr><tr><td> </td><td> </td><td> </td></tr><tr><td> </td><td> </td><td> </td></tr><tr><td> </td><td> </td><td> </td></tr></table>		4											<table border="1" style="display: inline-table;"><tr><td> </td><td>5</td><td> </td></tr><tr><td> </td><td> </td><td> </td></tr><tr><td> </td><td> </td><td> </td></tr><tr><td> </td><td> </td><td> </td></tr></table>		5											<table border="1" style="display: inline-table;"><tr><td> </td><td>8</td><td> </td></tr><tr><td> </td><td> </td><td> </td></tr><tr><td> </td><td> </td><td> </td></tr><tr><td> </td><td> </td><td> </td></tr></table>		8											<table border="1" style="display: inline-table;"><tr><td> </td><td>14</td><td> </td></tr><tr><td> </td><td> </td><td> </td></tr><tr><td> </td><td> </td><td> </td></tr><tr><td> </td><td> </td><td> </td></tr></table>		14											<table border="1" style="display: inline-table;"><tr><td> </td><td>15</td><td> </td></tr><tr><td> </td><td> </td><td> </td></tr><tr><td> </td><td> </td><td> </td></tr><tr><td> </td><td> </td><td> </td></tr></table>		15										
	1																																																																																											
	3																																																																																											
	4																																																																																											
	5																																																																																											
	8																																																																																											
	14																																																																																											
	15																																																																																											
<i>PRODUCTION FORECASTS</i>																																																																																												

4. DATA CHARACTERISTICS											
QUALIFIERS: YES = X					CRITICISMS: YES = X						
1	KEY DATA	☒	1	AVAILABILITY	☐						
2	COMPUTERIZED	☒	2	PERIODICITY	☐						
3		☐	3	DELAY	☒						
4		☐	4	CONTENT	☐						
5		☐	5	RELIABILITY	☐						
6		☐	6		☐						
7		☐	7		☐						
8		☐	8		☐						
9		☐	9		☐						
ORIGIN: <table border="1" style="display: inline-table;"><tr><td> </td><td> </td><td> </td><td> </td><td> </td><td> </td></tr></table> <i>PRODUCTION PLANNING</i>											
COMMENTS: <i>3. HAS TO BE AVAILABLE IN REAL TIME</i>											

Figure 6 Statistics by data

ISMOD-BAC • ANY CORPORATION • TOTAL SYSTEM • STA:D4												
• STATISTICS •												
DATA	Q.1 KEY	Q.2 COMP	C.1 AVLB	C.2 PRDC	C.3 DLAY	C.4 CONT	C.5 RELB	A	B	AB	RD	S
1111	110	0	20	0	45	95	6	120	115	235	56	.51
1112	47	2	2	0	0	0	0	94	2	96	0	.98
1113	11	0	0	0	0	16	0	15	16	31	8	.48
1121	0	0	2	0	0	3	3	1	5	6	4	.17
1122	26	16	4	4	19	25	11	22	36	58	22	.38
1123	29	36	0	11	16	3	8	33	21	54	8	.61
1131	80	163	3	0	52	122	1	38	133	171	103	.22
1132	12	23	0	0	0	0	0	23	0	23	0	1.00
1141	26	0	4	0	1	0	0	47	5	52	0	.90
1142	45	0	1	0	6	16	0	66	17	83	3	.80
1143	11	0	1	0	0	17	0	11	18	29	11	.38
1211	119	0	20	0	45	95	6	120	115	235	56	.51
1212	52	0	2	0	0	0	0	92	2	94	0	.98
1221	5	0	0	1	0	0	0	9	1	10	0	.90
1222	15	0	0	0	2	4	0	16	6	22	2	.73
1231	87	163	3	0	52	122	0	38	132	170	102	.22
1232	7	10	0	0	0	0	0	10	0	10	0	1.00
1241	20	0	4	0	1	0	0	47	5	52	0	.90
1242	39	0	1	0	6	16	0	66	17	83	3	.80
1243	15	0	1	0	0	17	0	11	18	29	11	.38
1311	119	0	20	0	45	95	6	120	115	235	56	.51
1312	45	0	2	0	0	0	0	92	2	94	0	.98
1321	2	0	0	0	1	0	0	4	1	5	0	.80
1322	8	0	0	0	1	0	0	10	1	11	0	.91
1323	7	17	0	0	1	0	0	16	1	17	0	.94
1331	72	163	3	0	52	122	0	38	132	170	102	.22
1332	4	10	0	0	0	0	0	10	0	10	0	1.00
1341	17	0	4	0	1	0	0	47	5	52	0	.90
1342	36	0	1	0	6	16	0	66	17	83	3	.80
1343	14	0	1	0	0	17	0	11	18	29	11	.38
1413	2	0	0	0	0	0	0	11	0	11	0	1.00
1422	4	0	0	0	0	0	0	9	0	9	0	1.00
1423	20	34	0	0	1	4	0	33	5	38	1	.87
2111	112	0	20	0	45	91	6	111	111	222	56	.50
2112	47	2	0	0	6	14	4	67	18	85	4	.79
2113	59	0	0	0	42	31	12	90	61	151	25	.60
2114	1	0	0	0	0	0	0	2	0	2	0	1.00
2121	8	5	0	0	0	4	0	17	4	21	1	.81
2122	10	16	0	0	0	16	16	1	16	17	15	.06

viewee works. Additionally, information is provided about the processes (tasks) that are performed by the individual and the event, if there is one, which causes the process to be executed. Once the processes are identified, the respondent indicates what data are used to perform each process and which process originates the data. During the diagnosis phase, this information allows the analyst to determine the data flow through the enterprise and the interdependency of processes. For each data item used, a respondent

is requested to answer some questions and criticisms defined by the study team.

For example, the questions presented may ask for such information as

- Are these data essential in performing the process?
- Are the data computerized?

Criticism may take the form of questions such as

- Are the data *available* when required?
- Do the data meet *frequency* requirements?
- Are the data available on a *delayed* basis, and not on time?
- Is the *content* reliable?

Diagnosis phase

Before defining the future information systems architecture, it is necessary to understand the current information system and its state of adequacy. The point is that the survey of the enterprise comprises all aspects of the information system, whether mechanized or manually supported. Thus, our analysis concerns itself with all the information processes in the business.

Understanding the current information system. The reports provided in this area are the Statistics Report and the Initial Matrix. We will first look at the Statistics Report, shown in Figure 6. It can be printed for each single dimension such as process, data, event, location, etc. Furthermore, it may be printed with up to six dimensions combined on one report. For example, Figure 7 shows the report printed by organization (hierarchy) within data view. The information displayed on the report shows the responses to the questions and criticisms requested during the survey. Additionally, we see the total usages (AB), the satisfactory (A) and the unsatisfactory (B) usages, as well as the overall percentage of satisfaction (S). The relative dissatisfaction (RD), computed as $RD = B \times (1 - S)$, shows the magnitude of dissatisfaction a particular item has in relation to others. Thus, data item 1121, in Figure 6, at 17 percent satisfaction is much less important than data item 1331 at 22 percent, since it has only 5 occurrences of dissatisfaction versus 132 occurrences for data item 1331.

Let us analyze some of the information on the report in Figure 6. The heading DATA may represent a data view, class, entity, or attribute, depending on how the study team defined data initially to the system. It can be seen that data item 1111 has a total number of 235 usages (AB) of which 110 (Q.1) are key. This means that the data are required to perform the process or that the process has to wait until the data are available. Note, however, that none of the data are computerized (Q.2) and also that 115 usages identified the data as unsatisfactory (B). When analyzing the unsatisfactory usages (C.1 to C.5), it should be noted that most criticism falls in the area of Content and Delay. If one or more criticisms are

marked by the same respondent for a particular data item, it is only counted once in the unsatisfactory usage column. The RD column shows that the relative dissatisfaction is 56, which puts this data item in the higher category of problem data.

Studying the report in Figure 7, we see that all Availability and Delay criticisms come from interviewee 3411, as well as most of the criticisms on Content. The RD column shows that user 3411 also accounts for the greatest contribution to dissatisfaction for these data.

On the basis of this analysis, the question that may be asked is, Why is this interviewee so different from the rest? What is the basic cause? How can this

The isolation and interaction matrix is the most heavily used during the ISMOD study.

problem be solved easily and cost-effectively? All that might be required in this case is some simple coordination between this user and other users of the data. In other situations a more comprehensive solution may need to be found.

The Initial Matrix (Figure 8) provides some of the same data found in the Statistics Report in a more global, enterprise-wide form. Again, any dimension can be shown on either axis of the matrix. Here, data groups 1 to 9 are compared to process group 1 to 8. Any dimension can have its detailed information combined into higher-level groups by use of a coding scheme which is not discussed here. Data items 1111 to 1423 in Figure 6 can be combined into a single data group 1 as shown in Figure 8.

The lines are labeled (A) Satisfactory, (B) Unsatisfactory, (AB) Total Usage, and (S) Satisfaction Percent. A total column and row are provided, representing the overall enterprise figures.

A closer look at the Initial Matrix shows that process group 4 has the lowest satisfaction of all process groups. Note that process group 4 is particularly

Figure 7 Statistics on two levels

ISMOD-BAC * ANY CORPORATION * TOTAL SYSTEM * STA:D4H4													

* STATISTICS *													

DATA	HIER	Q.1 KEY	Q.2 COMP	C.1 AVLB	C.2 PRDC	C.3 DLAY	C.4 CONT	C.5 RELB	A	B	AB	RD	S
1111	2100	11	0	0	0	0	6	6	18	6	24	2	.75
1111	2300	6	0	0	0	0	0	0	12	0	12	0	1.00
1111	2400	6	0	0	0	0	14	0	0	14	14	14	.00
1111	2500	12	0	0	0	0	0	0	18	0	18	0	1.00
1111	3110	1	0	0	0	0	4	0	0	4	4	4	.00
1111	3241	7	0	0	0	0	12	0	0	12	12	12	.00
1111	3271	14	0	0	0	0	0	0	24	0	24	0	1.00
1111	3411	27	0	20	0	45	45	0	0	65	65	65	.00
1111	3441	2	0	0	0	0	8	0	0	8	8	8	.00
1111	3451	5	0	0	0	0	0	0	9	0	9	0	1.00
1111	3491	7	0	0	0	0	0	0	12	0	12	0	1.00
1111	4121	6	0	0	0	0	0	0	14	0	14	0	1.00
1111	4212	2	0	0	0	0	2	0	0	2	2	2	.00
1111	4231	0	0	0	0	0	0	0	2	0	2	0	1.00
1111	4610	0	0	0	0	0	4	0	0	4	4	4	.00
1111	5110	0	0	0	0	0	0	0	1	0	1	0	1.00
1111	5210	4	0	0	0	0	0	0	6	0	6	0	1.00
1111	5310	0	0	0	0	0	0	0	4	0	4	0	1.00
1111	****	110	0	20	0	45	95	6	120	115	235	56	.51
1112	2100	10	0	0	0	0	0	0	18	0	18	0	1.00
1112	2300	1	0	0	0	0	0	0	2	0	2	0	1.00
1112	2400	12	0	0	0	0	0	0	24	0	24	0	1.00
1112	3241	2	0	0	0	0	0	0	6	0	6	0	1.00
1112	3271	11	0	0	0	0	0	0	22	0	22	0	1.00
1112	3491	2	0	0	0	0	0	0	5	0	5	0	1.00
1112	4121	8	2	0	0	0	0	0	16	0	16	0	1.00
1112	4610	1	0	2	0	0	0	0	0	2	2	2	.00
1112	5110	0	0	0	0	0	0	0	1	0	1	0	1.00
1112	****	47	2	2	0	0	0	0	94	2	96	0	.98
1113	3271	0	0	0	0	0	0	0	10	0	10	0	1.00
1113	3451	0	0	0	0	0	0	0	1	0	1	0	1.00
1113	4121	11	0	0	0	0	16	0	4	16	20	13	.20
1113	****	11	0	0	0	0	16	0	15	16	31	8	.48

dissatisfied with data groups 5 and 9. The bottom totals, which reflect the overall enterprise satisfaction, reflect the fact that the enterprise as a whole is very dissatisfied with these two data groups. Also note that process group 4 accounts for 233 of the total 446 usages of data group 5 and 539 of the total 1107 usages of data group 9. Thus, process group 4 places a heavy demand on the data from these data groups. The previous reports may be used to pursue further questions, such as the specific data involved, by listing data within a "using" process for all proc-

esses in process group 4. It is possible to determine which processes create this information by listing the data from data groups 5 and 9 within "originating" processes. As can be seen, the analysis of the current information system and the reasons for certain phenomena can be quite exhaustive and variable based upon the nature of the problem. It is impossible to explore all the variations in this paper.

Once an understanding of the current system (manual and mechanized) is gained, it becomes necessary

Figure 8 Initial matrix

ISMOD-BAC * ANY CORPORATION * TOTAL SYSTEM * MAT:Q1D1										

* INITIAL MATRIX *										

DATA		1	2	3	4	5	6	7	8	9
PROC										*
1	A	129	94	14	4	3	20	5	70	99
	B	57	68	24	10	13	0	4	7	51
	AB	186	162	38	14	16	20	9	77	150
	S	.69	.58	.37	.29	.19	1.00	.56	.91	.66
2	A	62	43	5	13	0	192	7	9	17
	B	12	16	0	2	23	53	0	0	3
	AB	74	59	5	15	23	245	7	9	20
	S	.84	.73	1.00	.87	1.00	.78	1.00	1.00	.85
3	A	88	66	4	8	0	139	27	113	63
	B	23	12	35	8	1	29	0	10	0
	AB	111	78	39	16	1	168	27	123	63
	S	.79	.85	.10	.50	.00	.83	1.00	.92	1.00
4	A	677	427	77	230	26	25	11	150	85
	B	546	370	172	302	207	11	7	230	454
	AB	1223	797	249	532	233	36	18	380	539
	S	.55	.54	.31	.43	.11	.69	.61	.39	.16
5	A	151	150	69	100	34	118	30	312	68
	B	149	186	69	110	98	115	27	220	105
	AB	300	336	138	210	132	233	57	552	173
	S	.50	.45	.50	.48	.26	.51	.53	.59	.39
6	A	172	121	331	44	6	19	3	166	67
	B	135	109	371	11	4	4	10	149	80
	AB	307	230	702	55	10	23	13	315	147
	S	.56	.53	.47	.80	.60	.83	.23	.53	.46
7	A	0	0	18	2	2	6	50	19	0
	B	3	0	16	0	0	0	37	39	0
	AB	3	0	34	2	2	6	87	58	0
	S	.00	-	.53	1.00	1.00	1.00	.57	.33	-
8	A	67	79	39	26	8	102	17	35	10
	B	36	43	60	43	21	19	10	42	5
	AB	103	122	99	69	29	121	27	77	15
	S	.65	.65	.39	.38	.28	.84	.63	.45	.67
*•	A	1346	980	557	427	79	621	150	874	409
	B	961	804	747	486	367	231	95	697	698
	AB	2307	1784	1304	913	446	852	245	1571	1107
	S	.58	.55	.43	.47	.18	.73	.61	.56	.37

to combine the current processes into process sets or *subsystems*. In more traditionally designed information systems, processes have been combined into systems based upon the organization in which they are being performed. Thus, we find the same processes being performed in different ways using differ-

ent data by different organizations. What needs to be accomplished is that all processes that share common data are performed in the same way by different organizational entities. To do this, the processes are evaluated to determine the extent of interdependence using the principle of isolation and interaction.

Isolation and interaction matrix

The isolation and interaction matrix is the one most heavily used during the ISMOD study (Figure 9). It is used to determine how processes should be combined into subsystems based upon the interaction of data. In order to determine interaction between processes, it is necessary to identify the degree to which a process can stand on its own or is dependent on data provided by another process. ISMOD uses the concept of isolation and interaction to determine this.

Isolation—Isolation is defined as the amount of data required by a process and created by that process. Thus, if a process creates all the data it uses, it is said to be 100 percent isolated. Therefore, it can stand on its own as a system. However, as a process becomes more and more dependent on other processes for data, it is said to be dependent. Therefore, it must be considered for implementation with other processes into a subsystem.

Interaction—Interaction simply represents the amount of data flowing from one process to the next.

Figure 9 Isolation matrix

ISMOD-BAC • ANY CORPORATION • TOTAL SYSTEM • ISO:Q101											

• ISOLATION RATIOS • DI - MINIMUM: .00											

ORIG		1	2	3	4	5	6	7	8	AB	I C 1
PROC											
1	AB	300	5	34	162	105	61	5	0	672	.446
	IM		.359	.353	.347	.385	.306	.432			1
	IE		.465	.569	.705	.623	.636	.538			4
	DI		.106	.216	.358	.238	.329	.106			.358
2	AB	115	105	108	6	112	4	7	0	457	.230
	IM	.359		.243	.320	.342	.249	.274			2
	IE	.465		.384	.332	.448	.256	.285			3
	DI	.106		.141	.013	.107	.007	.011			.141
3	AB	247	45	158	28	74	51	23	0	626	.252
	IM	.353	.243		.319	.340	.254	.282			3
	IE	.569	.384		.336	.421	.287	.318			1
	DI	.216	.141		.017	.081	.033	.035			.216
4	AB	1514	51	51	1322	815	236	18	0	4007	.330
	IM	.347	.320	.319		.342	.306	.332			4
	IE	.705	.332	.336		.524	.382	.338			1
	DI	.358	.013	.017		.181	.076	.005			.358
5	AB	558	162	147	295	772	121	56	0	2111	.366
	IM	.385	.342	.340	.342		.314	.367			5
	IE	.623	.448	.421	.524		.430	.393			1
	DI	.238	.107	.081	.181		.116	.026			.238
6	AB	754	12	29	205	331	458	13	0	1802	.254
	IM	.306	.249	.254	.306	.314		.266			6
	IE	.636	.256	.287	.382	.430		.281			1
	DI	.329	.007	.033	.076	.116		.015			.329
7	AB	87	0	6	5	5	16	73	0	192	.380
	IM	.432		.282	.332	.367	.266				7
	IE	.538		.318	.338	.393	.281				1
	DI	.106		.035	.005	.026	.015				.106
8	AB	110	65	76	105	195	85	26	0	662	.000
	IM	.225	.094	.123	.283	.278	.186	.085			8
	IE	.307	.152	.182	.306	.349	.220	.116			1
	DI	.082	.058	.059	.022	.070	.034	.030			.082

It may be said to be light or heavy. Light interaction means more isolation, whereas heavy interaction indicates more dependency.

Clustering—It is the objective of an ISMOD study to determine, on the basis of data interaction of the processes, which processes must be combined into the various information subsystems of the enterprise, and how those subsystems relate to one another. To do this, the study team uses a concept called clustering. Clustering refers to the activity of combining processes into subsystems. To accomplish this result, each process must be analyzed in relation to every other process through a concept called coupling.

Coupling—Coupling is the procedure of analyzing the relationships of a process to every other process in terms of data sharing. If the interaction is high, two processes can be looked upon as a couple and hence be made part of the same subsystem. If the interaction is low, the two would not be considered to be ideal for common grouping.

Report explanation. Figure 9 is an example of an isolation matrix produced by ISMOD. It shows the using processes as rows and the originating processes as columns. Using the AB row for process 6 as an example, we note that it creates 458 elements which it uses. Also, this is out of a total of 1802 used in that process. Thus, the isolation (I) of process 6 is 25.4 percent ($458/1802$). It uses 205 data types from process 4, 754 from process 1, etc. Notice that process 8 is completely dependent, since it creates no data of its own, and process 1 is the most isolated, with 44.6 percent. The row identifier AB stands for total data (satisfactory and unsatisfactory).

Referring to Figure 9, we now explain the meaning of the IM, IE, and DI row values.

IM (Isolation Mean)—The IM value represents the average isolation between two processes if they are combined. It is derived by combining their individual isolation factors divided by the new base of total data types used after coupling.

IE (Isolation Extended)—IE represents the isolation that would result through coupling when the data shared by the two processes are also considered in addition to their individual isolations. For example, if processes 6 and 4 are to be considered as one process, the isolation of each process (458 and 1322) must be combined to form the new isolation. However, because processes 4 and 6 are combined, we

must also include the data shared between the two processes in the new isolation value, namely, the 205 data types that 6 uses from 4 and the 236 data types that 4 uses from 6.

DI (Difference in Isolation)—DI represents the improvement gained by combining a particular process with another process. It is the difference between the Isolation Extended and the Isolation Mean, and thus represents the percent improvement which can be gained by combining the two processes. In other words, it represents the percent overlap or shared data between the two processes.

In Figure 9, process 6 coupled with process 4 will have a 7.6 percent improvement. However, when process 6 is coupled with process 1 the improvement is 32.9 percent. This is indeed the best couple for process 6 when all the other DIs for process 6 are considered. Figure 10 is a continuation of the isolation matrix and shows the best couple (C) for each process in descending order. Note that process 6 is best coupled with process 1, rather than with processes 5, 4, 3, 7, and 2 in descending order of desirability. Note also that even though process 4 is only the third choice in coupling with process 6, it is the first choice of process 1, and its own number-one choice for coupling is also process 1. This means that these two processes are highly interdependent. Since process 6 should be coupled with process 1 as its first choice and process 1 with process 6 as its second choice, it may still be conceivable that process 6 and process 4 will be in the same subsystem when other factors known to the study team are considered.

Clustering. In order to eliminate origin/usage intersections that may not be of interest, the user may print isolation matrices specifying a particular DI value to be used. For example, the user may specify that any DIs of five percent or less are not of significant interest and are not to be printed by the system. This case is shown in Figure 11. Note that if a process is coupled with itself, the improvement is always zero. The DIs on the diagonal are always zero because the matrix is made square by the system and the data show the improvement if a process is coupled with itself. On all intersections where the DI is five percent or less, only the actual data type counts are printed, but no IM, IE, or DI values. This makes it easier to identify which processes should be clustered into a subsystem based upon their interaction at a particular DI level.

Figure 12 shows two clustered subsystems. One is shown enclosed in asterisks and one in plus signs. It

Figure 10 Complementary isolation matrix

ISMOD-BAC • ANY CORPORATION • TOTAL SYSTEM • ISO:Q101									
	AB	I	C 1	C 2	C 3	C 4	C 5	C 6	C 7
1	672	.446	1 4 .358	1 6 .329	1 5 .238	1 3 .216	1 7 .106	1 2 .106	
2	457	.230	2 3 .141	2 5 .107	2 1 .106	2 4 .013	2 7 .011	2 6 .007	
3	626	.252	3 1 .216	3 2 .141	3 5 .081	3 7 .035	3 6 .033	3 4 .017	
4	4007	.330	4 1 .358	4 5 .181	4 6 .076	4 3 .017	4 2 .013	4 7 .005	
5	2111	.366	5 1 .238	5 4 .181	5 6 .116	5 2 .107	5 3 .081	5 7 .026	
6	1802	.254	6 1 .329	6 5 .116	6 4 .076	6 3 .033	6 7 .015	6 2 .007	
7	192	.380	7 1 .106	7 3 .035	7 5 .026	7 6 .015	7 4 .005		
8	662	.000	8 1 .082	8 5 .070	8 3 .059	8 2 .058	8 6 .034	8 7 .030	8 4 .022

can be seen that processes are printed at the third level of detail (three-digit process code), and not at the higher level as in the previous examples. On those figures the data were summarized at the highest level to allow the reader an easy overview of all data for discussion purposes. Here, a lower level of detail has been chosen to show clustering. Notice that the interaction improvement factor (DI) is much lower on this matrix since the data flow represents a much lower level. Therefore, a DI of three percent may seem low at first, but may actually represent a large factor at this level.

When analyzing the matrix and the DI factors, the user may see how the two subsystems were derived.

Of interest may be process 181. Note that its heavy interdependence with processes 121 and 151 put it into the asterisk subsystem. The interaction with process 191 of the plus-sign subsystem is so small that it cannot be considered as an essential part of that subsystem.

The determination as to subsystem assignment should not be made on the basis of the matrix alone, but should also require some understanding of the process and the actual data used to determine whether the coupling and clustering suggested by the matrix is sensible.

Usage of the isolation and interaction matrix. The isolation and interaction matrix is primarily used by

the designer to determine the composition of subsystems based upon data interaction. The use of data interaction in determining subsystems is of extreme importance, since the extent to which data are shared is the key to building integrated data systems. Implementation of a shared data subsystem can be seriously affected if the clustering decision was not based upon data usage and sharing.

The user may experiment with different DI values which will allow easier identification of the subsys-

tems, but should always apply human knowledge to the data provided by ISMOD.

Creating subsystems

After organizing the current processes into their respective subsystems, the analyst can perform an analysis on the proposed subsystem. For instance, in printing the Statistics Report and the Initial Matrix, the current satisfaction can be determined by subsystem. Strong or weak subsystems can be identified. Furthermore, printing the Statistics Report by orig-

Figure 11 Isolation matrix with other DI minimum

ISMOD-BAC • ANY CORPORATION • TOTAL SYSTEM • ISO:Q101									

• ISOLATION RATIOS •									

ORIG		1	2	3	4	5	6	7	8
PROC									
1	AB	300	5	34	162	105	61	5	0
	IM		.359	.353	.347	.385	.306	.432	
	IE		.465	.569	.705	.623	.636	.538	
	DI		.106	.216	.358	.238	.329	.106	
2	AB	115	105	108	6	112	4	7	0
	IM		.359	.243		.342			
	IE		.465	.384		.448			
	DI		.106	.141		.107			
3	AB	247	45	158	28	74	51	23	0
	IM		.353	.243		.340			
	IE		.569	.384		.421			
	DI		.216	.141		.081			
4	AB	1514	51	51	1322	815	236	18	0
	IM		.347			.342	.306		
	IE		.705			.524	.382		
	DI		.358			.181	.076		
5	AB	558	162	147	295	772	121	56	0
	IM		.385	.342	.340	.342	.314		
	IE		.623	.448	.421	.524	.430		
	DI		.238	.107	.081	.181	.116		
6	AB	754	12	29	205	331	458	13	0
	IM		.306		.306	.314			
	IE		.636		.382	.430			
	DI		.329		.076	.116			
7	AB	87	0	6	5	5	16	73	0
	IM		.432						
	IE		.538						
	DI		.106						
8	AB	110	65	76	105	195	85	26	0
	IM		.225	.094	.123	.278			
	IE		.307	.152	.182	.349			
	DI		.082	.058	.059	.070			

Figure 12 Clustering into subsystems

ISMOD-BAC * ANY CORPORATION * CLUSTERING * ISO:Q303												

* ISOLATION RATIOS * DI - MINIMUM: .03												

ORIG		111	121	131	141	151	161	171	181	191	211	212 ...
PROC												
111	AB	3	3	0	0	2	0	0	0	0	0	0
	DI	.045										
121	AB	4	7	0	0	3	0	0	0	1	0	0
	DI	.045				.056						
131	AB	4	3	0	3	3	0	0	0	1	0	0
	DI	.034			.058	.081						
141	AB	2	2	0	2	2	0	0	0	1	0	0
	DI				.118							
151	AB	0	0	0	0	0	0	0	0	0	0	0
	DI											
161	AB	5	6	0	0	3	72	35	0	1	0	0
	DI						.226					
171	AB	5	3	0	0	3	47	35	0	0	0	0
	DI						.226					
181	AB	3	4	0	0	3	0	0	0	1	0	0
	DI	.036				.052						
191	AB	0	0	0	0	0	0	0	0	18	0	0
	DI											
211	AB	0	0	0	1	0	0	7	0	4	2	0
	DI						.034		.042			
212	AB	0	0	0	1	0	0	6	0	4	0	0
	DI						.035		.063			

inating processes within a subsystem, one can determine sources of information problems.

Also, new processes or systems can be added to the existing enterprise information model. It should be

noted that this is rarely the case in real life, since the enterprise processes are quite well defined, and that unless the enterprise enters into some new type of activity, no changes in the functioning enterprise processes are required.

Figure 13 Simple simulation matrix

ISMOD-BAC * ANY CORPORATION * TOTAL SYSTEM * SIM:QID1											

* SIMPLE SIMULATION *											

PERFECT DATA: 5											
DATA											
PROC		1	2	3	4	5	6	7	8	9	*
1	A	129	94	14	4	16	20	5	70	99	451
	B	57	68	24	10	0	0	4	7	51	221
	AB	186	162	38	14	16	20	9	77	150	672
	S	.69	.58	.37	.29	1.00	1.00	.56	.91	.66	.67
2	A	62	43	5	13	23	192	7	9	17	371
	B	12	16	0	2	0	53	0	0	3	86
	AB	74	59	5	15	23	245	7	9	20	457
	S	.84	.73	1.00	.87	1.00	.78	1.00	1.00	.85	.81
3	A	88	66	4	8	1	139	27	113	63	509
	B	23	12	35	8	0	29	0	10	0	117
	AB	111	78	39	16	1	168	27	123	63	626
	S	.79	.85	.10	.50	1.00	.83	1.00	.92	1.00	.81
4	A	677	427	77	230	233	25	11	150	85	1915
	B	546	370	172	302	0	11	7	230	454	2092
	AB	1223	797	249	532	233	36	18	380	539	4007
	S	.55	.54	.31	.43	1.00	.69	.61	.39	.16	.48
5	A	151	150	69	100	132	118	30	312	68	1130
	B	149	186	69	110	0	115	27	220	105	981
	AB	300	336	138	210	132	233	57	532	173	2111
	S	.50	.45	.50	.48	1.00	.51	.53	.59	.39	.54
6	A	172	121	331	44	10	19	3	166	67	933
	B	135	109	371	11	0	4	10	149	80	869
	AB	307	230	702	55	10	23	13	315	147	1802
	S	.56	.53	.47	.80	1.00	.83	.23	.53	.46	.52
7	A	0	0	18	2	2	6	50	19	0	97
	B	3	0	16	0	0	0	37	39	0	95
	AB	3	0	34	2	2	6	87	58	0	192
	S	.00	-	.53	1.00	1.00	1.00	.57	.33	-	.51
8	A	67	79	39	26	29	102	17	35	10	404
	B	36	43	60	43	0	19	10	42	5	258
	AB	103	122	99	69	29	121	27	77	15	662
	S	.65	.65	.39	.38	1.00	.84	.63	.45	.67	.61
*	A	1346	980	557	427	446	621	150	874	409	5810
	B	961	804	747	486	0	231	95	697	698	4719
	AB	2307	1784	1304	913	446	852	245	1571	1107	10529
	S	.58	.55	.43	.47	1.00	.73	.61	.56	.37	.55

The question remains, Where is the place to start expending the limited resources of the enterprise in order to improve the overall quality of the currently proposed set of subsystems? We know, by interrogating ISMOD reports, where information quality

weaknesses occur. It is also known where the sources of those weaknesses are. What we do not know is what the impact would be of expending effort in a particular area of the enterprise. As we have seen, many of the processes and subsystems are interre-

Figure 14 Total simulation matrix

ISMOD-BAC * ANY CORPORATION * TOTAL SYSTEM * TOT:Q1D1											
***** * TOTAL SIMULATION * *****											
DATA											
PROC		*	1	2	3	4	5	6	7	8	9
1	A	438	495	506	462	448	451	438	442	445	489
	B	234	177	166	210	224	221	234	230	227	183
	AB	672	672	672	672	672	672	672	672	672	672
	S	.65	.74	.75	.69	.67	.67	.65	.66	.66	.73
2	A	348	360	364	348	350	371	401	348	348	351
	B	109	97	93	109	107	86	56	109	109	106
	AB	457	457	457	457	457	457	457	457	457	457
	S	.76	.79	.80	.76	.77	.81	.88	.76	.76	.77
3	A	508	531	520	543	516	509	537	508	518	508
	B	118	95	106	83	110	117	89	118	108	118
	AB	626	626	626	626	626	626	626	626	626	626
	S	.81	.85	.83	.87	.82	.81	.86	.81	.83	.81
4	A	1708	2254	2078	1880	2010	1915	1719	1715	1938	2162
	B	2299	1753	1929	2127	1997	2092	2288	2292	2069	1845
	AB	4007	4007	4007	4007	4007	4007	4007	4007	4007	4007
	S	.43	.56	.52	.47	.50	.48	.43	.43	.48	.54
5	A	1032	1181	1218	1101	1142	1130	1147	1059	1252	1137
	B	1079	930	893	1010	969	981	964	1052	859	974
	AB	2111	2111	2111	2111	2111	2111	2111	2111	2111	2111
	S	.49	.56	.58	.52	.54	.54	.54	.50	.59	.54
6	A	929	1064	1038	1300	940	933	933	939	1078	1009
	B	873	738	764	502	862	869	869	863	724	793
	AB	1802	1802	1802	1802	1802	1802	1802	1802	1802	1802
	S	.52	.59	.58	.72	.52	.52	.52	.52	.60	.56
7	A	97	100	97	113	97	97	97	134	136	97
	B	95	92	95	79	95	95	95	58	56	95
	AB	192	192	192	192	192	192	192	192	192	192
	S	.51	.52	.51	.59	.51	.51	.51	.70	.71	.51
8	A	383	419	426	443	426	404	402	393	425	388
	B	279	243	236	219	236	258	260	269	237	274
	AB	662	662	662	662	662	662	662	662	662	662
	S	.58	.63	.64	.67	.64	.61	.61	.59	.64	.59
*	A	5443	6404	6247	6190	5929	5810	5674	5538	6140	6141
	B	5086	4125	4282	4339	4600	4719	4855	4991	4389	4388
	AB	10529	10529	10529	10529	10529	10529	10529	10529	10529	10529
	S	.52	.61	.59	.59	.56	.55	.54	.53	.58	.58

lated to varying degrees. Thus, if a particular area is selected for rework and improvement, what will be the effect on all systems of the enterprise that have an interest in the data created and shared by the selected subsystem? The simulation facility is used to answer this question.

Simulation

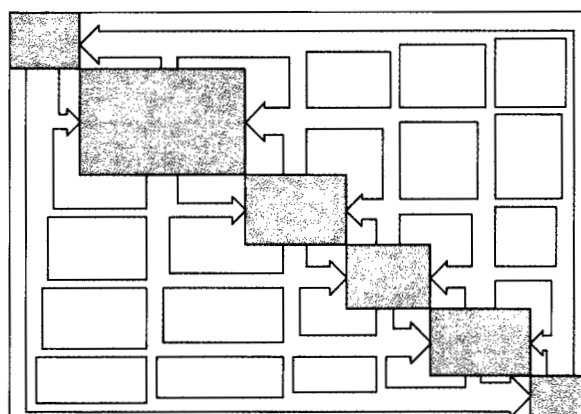
Simulation gives the architect the ability to formulate an implementation strategy and determine its value to the enterprise in terms of information satisfaction. Figure 13 is an example where the designer specified

Figure 15 Simulations on the subsystems

SUB-SYSTEM	PRESENT	SIMULATIONS (DATA CREATED IN SUBSYSTEMS)					
	S%	1	2	3	4	5	6
1	64		/	+9	/	+6	+2
2	46	+2		+8	+22	/	/
3	43	/	+9		/	+2	/
4	72	/	+11	+1		/	/
5	65	+4	/		+2	+8	+2
6	90	/	/	/	/	/	
TOT	57	+3	+9		+7	+3	+1



Figure 16 Flow of data between various processes



that data type 5 is assumed to have perfect satisfaction. Thus, we see that for each process the system has forced the satisfaction levels for data type 5 to 100 percent. The total column shows the effect on each process and the enterprise if this particular data type were to have perfect satisfaction. One way this might be achieved is by developing a data base

system for the subsystem which creates and has responsibility for a specific data type.

Only dimensions in the column positions of the matrix can be simulated. The system prints a summary matrix (Figure 14) showing the effect of each column dimension being made perfect. Figure 14 shows the original level of satisfaction based on the interviews in the asterisk column. All of the other columns show the result of making each data type perfect individually and separate from the other data types. So, if data type 1 is assumed to be perfect, the overall enterprise satisfaction could be improved from 52 percent to 61 percent. For data type 2 it would be 59 percent, and so on. Clearly, it may be seen that data type 1 shows the greatest potential for improvement.

The simulation facility is useful in determining where the greatest return in terms of satisfaction can be achieved by implementing a particular data system. Figure 15 shows a comparison of the improvement that could be achieved across subsystems if each subsystem had its data made perfect.

Analyzing the matrix, we see that, for example, subsystem 1 currently has a satisfaction level of 64 percent. However, if its own data were made perfect, this satisfaction level could be raised by 18 percent. If the data of subsystem 3 were made perfect, the satisfaction of subsystem 1 would only rise 9 percent. Thus, we can draw the conclusion that our best option for implementation is subsystem 1. Looking at the total chart, we see that almost all subsystems would like to have a project for themselves, except for subsystem 5. This system tells us that an effort in subsystem 3 would be more beneficial than a project for itself. The total line shows a weighted average of all of the improvement in satisfaction across all subsystems. Thus, it appears that subsystem 3 holds the greatest potential and payoff for the enterprise in terms of a new information system development project.

The reader should note carefully that this decision was based on information quality issues by considering data as a shared enterprise resource. It was not based upon the most insistent user department, strongest internal politician, or best guess by the data processing department.

Now that we have isolated the various subsystems and understand their interaction based on data sharing, we may proceed to continue our architecture

Figure 17 Theoretical responsibilities of managers

BUSINESS PROCESSES		INTERVIEWEES									
		GEN MGT	FINANCE		PLANT A		PLANT B			MKTG	
		1	2.1	2.2	4.1	4.2	5.1	5.2	5.3	6.1.1	6.1.2
PLANS & CONTROLS	1	▶	●	●	●		●			●	●
COMMERCIAL	2	▶	●							◐	◐
DISTRIBUTION	3		●		▶	◐	▶	◐			
PRODUCTION	4		●		▶	◐	▶		◐		
PERSONNEL	5	▶	●					●			
FINANCE	6	▶	◐	●	●		●			●	●
OTHERS	9		●		●				●		

▶ FINAL DECISION FOR PROCESS
● RESPONSIBLE FOR PROCESS
◐ INVOLVED IN PROCESS

approach with the chosen subsystem. That is to say, we use the same procedure again for the selected subsystem. We determine its subsystems, simulate, find the appropriate (sub)-subsystem, etc., until a level of detail is reached where individual processes can be identified for which procedures can be defined to be performed by an individual or create a specification to be programmed by a data processing professional.

Figure 16 represents the flow of data between each of the various processes. Each subsystem may have its own data-flow (isolation) matrix, resulting in a completely decomposed architecture.

Validation

Once the enterprise has implemented the proposed subsystem and the organization has had a reasonable amount of time to adjust to the new procedures, a new survey on the implemented part of the subsystem can be performed, and results can be compared to determine how well the original objectives were achieved. It may be decided that because of the

implementation of the first subsystem our initial prioritization of subsystems needs to be resimulated on the basis of the new satisfaction statistics available.

Other uses of the system

Organizations which have performed studies in the past have found many additional uses for the data provided by the system. Figure 17 is a matrix prepared by the study team during the Company Analysis Phase and shows the theoretical responsibilities of the various managers in the organization.

As can be seen, Figure 18 represents the actual involvement based on the interviews. Theory and reality do not necessarily coincide. For example, in Figure 17 plant interviewee 4.1 is shown to be the final decision authority (marked with the triangle) for process 3 (Distribution) and has no involvement of any kind in process 2 (Commercial). Yet, when comparing interviewees to the processes and looking at their actual data usages, it is apparent that interviewee 4.1 uses only one data item for Distribution.

Figure 18 Actual involvement matrix

BUSINESS PROCESSES		INTERVIEWEES										
		GEN MGT	FINANCE		PLANT A		PLANT B			MKTG		TOTAL
		1	2.1	2.2	4.1	4.2	5.1	5.2	5.3	6.1.1	6.1.2	
PLANS & CONTROLS	1	9	0	5	226	7	1	1	2	23	1	436
COMMERCIAL	2	27	0	16	169	15	0	27	6	253	235	1161
DISTRIBUTION	3	0	0	0	1	15	0	7	0	0	18	35
PRODUCTION	4	21	0	25	283	159	78	188	29	1	15	1122
PERSONNEL	5	8	1	7	34	22	5	16	9	10	16	235
FINANCE	6	8	31	38	0	3	2	3	3	0	1	139
OTHERS	9	3	0	2	43	7	2	3	4	9	6	141
TOTAL		76	32	93	756	228	88	245	53	296	292	

Also, he is heavily involved in the Commercial process with 169 usages, when he was supposed to have no involvement at all.

The system allows the user to compare any dimension against any other dimension and will produce information which is part of that relationship. Whether that information is meaningful must be decided by the user. Thus, one can compare the amount of data used by process/location, process/event, data/organization, etc.

In its present state, the system assists in process, or what is sometimes called function, architecture development. However, some basic information may be available in the system to assist with preliminary data and network architecture efforts.

Summary

ISMOD was developed to help enterprises design the architecture of their information systems based on the concept of data as an enterprise resource. The

key word here is architecture. Since the system allows the enterprise to keep its information architecture in mechanized form, it provides for continuous use of the information system model during the various levels of strategic planning, functional design, and operational implementation. Being able to validate the success of a newly implemented system and comparing it to the preimplementation environment is an important feedback mechanism.

Acknowledgments

I would like to express appreciation to Michel Veys of IBM Belgium for his imagination, talent, and many hours of hard work in creating ISMOD. He has been patient with many of us who have placed new system requirements on him and are continuing to do so. Also, I would like to thank Dr. Herwig Komarek of IBM Germany for his very extensive insight and continual research into the field of enterprise information architecture. His influence and work have been very helpful in our effort of bringing some semblance of order to the difficult task of integrated data system development.

Bibliography

1. R. Ambrosetti, T. A. Cirianni, and R. Pennacchi, "An application analyzer," *IBM Systems Journal* 23, No. 4, 336-350 (1984).
2. *Business Systems Planning—Information Systems Planning Guide*, Application Manual, GE20-0527, IBM Corporation; available through IBM branch offices.
3. P. P. Chen, *Entity-Relationship Approach to Systems Analysis and Design*, UCLA, Los Angeles, CA (December 1979).
4. *Data Architecture*, Part of a set of manuals published by IBM Information Systems Management European Marketing Investment Project (EuroMIP) (IBM Eurocoordination), under the name Information Systems Management; available from IBM Eurocoordination or IBM Europe marketing organizations.
5. G. B. Davis, "Strategies for information requirements determination," *IBM Systems Journal* 21, No. 1, 4-30 (1982).
6. C. Gane and T. Sarson, *Structured Systems Analysis: Tools and Techniques*, Prentice-Hall, Inc., Englewood Cliffs, NJ (1979).
7. K. P. Hein and M. Veys, *Information System Model and Architecture Generator—Study Guide*, Program Offering Manual, SH20-6651, IBM Corporation (December 1984); available through IBM branch offices.
8. H. Komarek, *KSS bei der ESSO A.G.* (in German). Example of the ESSO study, No. 223-83, IBM Germany, Stuttgart (1983).
9. R. L. Nolan, "Managing the crisis in data processing," *Harvard Business Review* 57, No. 2, 115-126 (March-April 1979).
10. R. L. Nolan, "Restructuring the data processing organization for data resource management," *Information Processing 77*, North-Holland Publishing Co., Amsterdam, The Netherlands (1977), pp. 261-265.
11. M. M. Parker, "Enterprise information analysis: Cost-benefit analysis and the data-managed system," *IBM Systems Journal* 21, No. 1, 108-123 (1982).
12. P. Pellaumail, *La Méthode AXIAL—Analyse et Conception de Système d'Information Assistées par Logiciels* (in French); a set of manuals, Z4F2-0020/0025, IBM France, Paris (1983-84).
13. C. Rochat, "Planification d'un Système d'Information" (in French). Article in *Bulletin IBM*, Volume 131, the regular publication to customers of IBM Switzerland, Geneva (May 1982).
14. J. G. Sakamoto, *Use of DB/DC Dictionary to Support Business Systems Planning Studies: An Approach*, G320-2705, IBM Corporation, Los Angeles Scientific Center (July 1980); available through IBM branch offices.
15. M. Vetter, "Aufbau betrieblicher Informationssysteme" (in German), published as part of a series called *Leitfaden der angewandten Informatik* by B. G. Teubner, Stuttgart, Germany.
16. M. Vetter, *Database Design Methodology*, Prentice-Hall International, Englewood Cliffs, NJ (1981).
17. M. M. Veys, *Information System Model and Architecture Generator—System Guide*, Program Offering Manual, LY20-0975, IBM Corporation (December 1984); available with purchase of product.
18. M. M. Veys, "ISS pour une logique de la productivité" (in French). Article in *Information*, the regular publication to customers of IBM Belgium, Volume 90, Page 2 (Brussels, 1Q-1979). Also published in Dutch with the same references in *Berichten*.
19. P. D. Walker, "Next in MIS: 'Data managed' systems design," *Computer Decisions* 1, No. 12 (December 1969).
20. J. A. Zachman, "Business System Planning and Business Information Control Study: A comparison," *IBM Systems Journal* 21, No. 1, 31-53 (1982).

K. Peter Hein IBM National Accounts Division, P.O. Box 45080, Salt Lake City, Utah 84145. Mr. Hein is a Senior Systems Engineer with many years of experience in Business Systems Planning, top-down system architecture, and large-scale system integration. He has lectured extensively at conferences and seminars on these subjects in the U.S. and abroad. At present, he leads an advanced technology project which deals with the automation of a complete methodology from the Strategic Information System Plan (BSP) to application creation with the use of professionals who are not trained in data processing. Mr. Hein graduated from the University of Utah with a B.S. and an M.B.A. degree. He is also a graduate of the IBM Systems Research Institute, the U.S. Army Command and General Staff College, and the U.S. Army War College.

Reprint Order No. G321-5250.